Group Contribution Based Estimation of Pure Component Properties

Jorge Marrero
Center of Pharmaceutical Chemistry, POB 16042, Havana, Cuba
Rafiqul Gani*
CAPEC, Department of Chemical Engineering, Technical University of Denmark
DK-2800 Lyngby, Denmark
*Corresponding author <rag@kt.dtu.dk>

Abstract

A new method for the estimation of properties of pure organic compounds is presented. Estimation is performed at three levels. The primary level uses contributions from simple groups that allow describing a wide variety of organic compounds while the higher levels involves polyfunctional and structural groups that provide more information about molecular fragments whose description through first-order groups is not possible. The presented method allows estimations of the following properties: normal boiling point, critical temperature, critical pressure, critical volume, standard enthalpy of formation, standard enthalpy of vaporization, standard Gibbs energy, normal melting point and standard enthalpy of fusion. The group contribution tables have been developed from regression using a data set of more than 2000 compounds ranging from C=3 to C=60, including large and complex polycyclic compounds. Compared to the currently used group-contribution methods, the new method makes significant improvements both in accuracy and applicability.

Keywords: Group Contribution; Property Prediction

Introduction

The basis for the design and simulation of many chemical processing units is a set of physical and thermodynamic properties of compounds in the process that undergo some form of transformation. It is not always possible, however, to find experimental values of properties for the compounds of interest in the literature. Since it is not practical either to measure them as the need arises, estimation methods are generally employed in this and other similar situations.

For the estimation of properties of pure compounds, group-contribution methods, Joback and Reid [1], Lydersen [2], Ambrose [3], Klincewicz and Reid [4], Lyman et al. [5], Horvath [6], have been widely used. In these methods, the property of a compound is a function of structurally dependent parameters, which are determined by summing the number frequency of each group occurring in the molecule times its contribution. These methods provide the advantage of quick estimates without requiring substantial computational resources. Many of these methods are, however, of questionable accuracy, unable to distinguish among isomers and have limited applicability due to the oversimplification of the molecular structure representation as a result of the use of a simple group-contribution approach and relatively small data set used for estimation of group contributions.

To overcome these limitations, several attempts have been reported in the literature. Constantinou et al. [7, 8] have proposed a quite complex estimation technique, which is based on conjugate forms (alternative formal arrangements of valence electrons). This

technique provides accurate estimations of several properties of pure compounds and allows capturing the differences among isomers. However, the generation of conjugate forms is a nontrivial issue and requires a symbolic computing environment, Prickett et al. [9]. A less complex method has been proposed by Constantinou and Gani [10], which performs the estimation at two levels: the basic level uses contributions from first-order simple groups, while the second level uses a small set of second-order groups having the first-order groups as building blocks. The role of the second-order groups is to consider, in some extent, the proximity effects and to distinguish among isomers. Marrero-Morejón and Pardillo-Fontdevila [11] proposed another technique that considers the contributions of interactions between bonding groups instead of the contributions of simple groups, which allows the distinction of a large number of isomers.

Despite the advantages of above-mentioned methods, however, their ranges of applicability are still quite restricted. Properties of large, complex and polyfunctional substances, of interest in biochemical and environmental studies, cannot be accurately estimated by using the current available methods. Due to the relatively small data sets used in the development of these methods, which usually includes just a few hundred of relatively simple compounds, the predictive capability usually breaks down when dealing with large, polycyclic or polyfunctional molecules. Also, most of the existing group-contribution techniques do not include suitable groups for representing complex molecules such as the ones of biochemical or environmental importance. Motivated by these drawbacks, our efforts have been focused on developing a new group-contribution

method that allows more accurate and reliable estimations of a wide range of chemical substances including large and complex compounds.

In our method, the estimation is performed at three levels. The basic level has a large set of simple groups that allow describing a wide variety of organic compounds. However, these groups capture only partially the proximity effects and are unable to distinguish among isomers. For this reason, the first level of estimation is intended to deal with simple and monofunctional compounds. The second level involves groups that permit a better description of proximity effects and differentiation among isomers. The second level of estimation is consequently intended to deal with polyfunctional, polar or non-polar, compounds of medium size, C=3 to C=6, and aromatic or cycloaliphatic compounds with only one ring and several substituents. The third level has groups that provide more structural information about molecular fragments of compounds whose description is insufficient through the first and second level's groups. The third level of estimation allows estimation of complex heterocyclic and large (C=7 to C=60) polyfunctional acyclic compounds. The ultimate objective of the proposed multi-level scheme is to enhance the accuracy, reliability and the range of application for a number of important pure component properties.

Development of the New Method

In the new method, the molecular structure of a compound is considered to be a collection of three types of groups: first-order groups, second-order groups and third-

order groups. The representation of a given compound through these groups is based on the following set of groups.

1. In the first level, groups describing the entire molecule must be selected. For example, CH3COCH2COCH(CH3)CH3 is described in the following way: (1) CH3CO, (1) CH2CO, (2) CH3, (1) CH2. In the case of aromatic substituents, groups of type aC-R must be chosen. For example, acetophenone is described by (1) aC-CO, (5) aCH and (1) CH3. The same molecular fragment can not be represented by more than one group. For example, trimethylurea is represented by (1) NHCON, (3) CH3. To use groups CH3NH or CH3N would be wrong because the nitrogen atoms would be covered more than one time.

2. In the second and third levels, the entire molecule does not need to be described by groups and the same molecular portion can be covered by more than one group. For example, cyclohexanol has only CHcyc-OH as a second-order group and cyclohexylmethacrylate is represented by the second-order groups CHcyc-OOC, CHn=CHm-COO and CH3-CHn=CHm. Contrary to the case of first-order level, there can be molecules that do not need any second-order or third-order groups (eg. acetophenone). There can be compounds that do not need any second-order group but need third-order groups such as for diphenyl sulfide in where a third-order group aC–S–aC is needed.

The property-estimation model has the form of the following equation.

$$f(X) = \sum_i N_i C_i + w \sum_j M_j D_j + z \sum_k O_k E_k \qquad (1)$$

In Eq. (1), C_i is the contribution of the first-order group of type-i that occurs N_i times, D_j is the contribution of the second-order group of type-j that occurs M_j times and E_k is the contribution of the third-order group of type-k that has O_k occurrences in a compound. In the first level of estimation, the constants *w* and *z* are assigned zero values because only first-order groups are employed. In the second level, the constants *w* and *z* are assigned unity and zero values respectively because only first and second-order groups are involved while in the third level, both *w* and *z* are set to unity values. The left-hand side of Eq. (1) is a simple function *f(X)* of the target property *X*. The selection of this function has been based on the following criteria:

1. The function has to achieve additivity in the contributions C_i, D_j and O_k.
2. It has to exhibit the best possible fit of the experimental data.
3. It should provide a good extrapolating capability and therefore a wide range of applicability.

According to these criteria, the selected functions are the same as used by Constantinou and Gani [10]. The target properties as well as their corresponding estimation functions are listed in Table 1. The symbols *Tm1i, Tb1i, Tc1i, Pc1i, Vc1i, Gf1i, Hf1i, Hv1i, Hfus1i* represent the contributions (C_i) of the first-order groups for the correspondig properties. Similarly, *Tm2j, Tb2j, Tc2j, Pc2j, Vc2j, Gf2j, Hf2j, Hv2j, Hfus2j* and *Tm3k, Tb3k, Tc3k, Pc3k, Vc3k, Gf3k, Hf3k, Hfus3k* represent the contributions (D_j) and (O_k.) of the second and third-order groups, respectively. The *Tm0, Tb0, Tc0, Pc1, Pc2, Vc0, Gf0, Hf0, Hv0,*

Hfus0 are additional adjustable parameters of the estimation models or universal constants.

The determination of the adjustable parameters of the models, that is, the contributions C_i, D_j and O_k, has been divided into a three-step regression procedure:

1. Regression is carried out to determine the contributions (C_i) of the first-order groups and the universal constants of the models while w and z are set to zero.
2. Then, w is set to unity, z is set to zero and another regression is performed using the C_is and universal constants calculated in the previous step to determine the contributions (D_j) of the second-order groups.
3. Finally, both w and z are assigned to unity and, using the universal constants of the models, C_is and D_js obtained as results of the previous steps, the contributions (O_k.) of the third-order groups are determined.

This stepped regression scheme ensures the independence among contributions of first, second and third order. Besides, the contributions of the higher levels act as corrections to the approximations of the lower levels. The total of the contributions C_i, D_j and O_k for the nine properties listed earlier can be obtained from the authors. The universal constants determined in the first step of the regression scheme are listed in Table 2. The optimization algorithm used for the data fitting was the Levenberg-Marquardt technique and the objective function was to minimize the sum of squares of the differences between experimental and estimated values of the target properties. The experimental data used in regression has been obtained from a comprehensive data bank of property

values developed at CAPEC-DTU [12] through a systematic search of several data sources. Property values have been included in this collection after a rigorous analysis of their reliability.

Results and Discussion

Table 3 presents for each property the standard deviation, the average absolute error and the average relative error for the first, second and third approximations. The number of experimental values used in the first regression step is also given. The statistics offered for the second and third approximations encompass all the data points, even those corresponding to compounds in which no second order or third order groups occur (and consequently not used in the second and third regression steps). Therefore, the average deviations given for the third approximation characterize the global results of the three subsequent approximations. Furthermore, due to the low number of available experimental values of enthalpies of vaporization at 298 K for complex and heterocyclic compounds, the contributions of third-order groups to this property have not been considered in this paper. A comparison of the average deviations obtained as results of the second and third regression steps is shown in Table 4, which does include the actual number of data points used in each step, that is, the number of compounds in which second and third order groups occur. For each set of compounds, the average deviations corresponding to both the current and previous step are presented in order to illustrate the improvement in accuracy achieved in each step.

The reliability of the estimation equations obtained from the regression steps has been tested for each property by performing a least-square analysis in which a randomly

conformed subset of the *N* experimental data points has been excluded from the full data set. Then, the mean-square residual *J* defined as follows,

$$J = \sqrt{\frac{\sum (X_i - Y_i)^2}{N}} \qquad (2)$$

was calculated. In Eq. 2, *N* is the number of data points excluded from the full data set, X_i is the property value of the compound *i* estimated by the full regression, and Y_i is the property value of the same compound estimated by the partial regression. For all the properties, the residuals are smaller than the estimation errors reported in Table 3, confirming the reliability of the method.

A fair comparison with other existing group-contribution methods is impossible since no other method exhibit the wide-ranging applicability of the proposed method. Moreover, the new method is able to deal with classes of compounds that cannot be handled by other widely used methods. The reason is that, compared to other methods, a significantly larger data set has been used in the development of the new method as well as a larger and comprehensive set of groups. However, in order to make a quite acceptable comparison of the proposed method with another classical group-contribution method, we have calculated the contributions of the groups used by Joback and Reid [1] and recalculated our group contributions using the estimation models reported in Table 1 and a common set of compounds that can be described by both group schemes. A comparison between the results obtained from both methods, the modified Joback's and the proposed one after the third approximation, is presented in Table 5. Clearly, the new method exhibits a much better accuracy.

Conclusion

The application of three different sets of functional groups, one for a first-order approximation and two successive ones for refining the estimations for complex, large and heterocyclic compounds has led to a new group contribution method for the estimation of important physical and thermodynamic properties. Compared to other currently used estimation methods, the proposed method exhibits an improved accuracy and a considerably wider range of applicability to deal with chemical, biochemical and environmental-related compounds. Even for lower molecular weight organic compounds, the larger set of first-order groups provides not only a wider range of application but also an improved accuracy. A computer program is also being developed for automatic selection of groups.

References

[1] K. G. Joback and R. C. Reid, “Estimation of Pure-Component Properties from Group-Contributions”, Chem. Eng. Comm., 57, 233 (1987).

[2] A. L. Lydersen, “Estimation of Critical Properties of Organic Compounds”, Univ. Cool. Exp. Stn., Rept., Madison, WI, April (1955).

[3] D. Ambrose, “Correlation and Estimation of Vapor-Liquid Critical Properties: I. Critical Temperatures of Organic Compounds”, Nat. Physical Lab., Teddington, UK, NPL Rep. Chem., 92 (1978).

[4] K. M. Klincewicz and R. C. Reid, “Estimation of Critical Properties with Group-Contribution Methods”, AIChE J., 30, 137 (1984).

[5] W. J. Lyman, W. F. Reehl and D. H. Rosenblatt, “Handbook of Chemical Property Estimation Methods”, Americal Chemical Soc., Washington, DC (1990).

[6] A. L. Horvath, “Molecular Design”, Elsevier, Amsterdam (1992).

[7] L. Constantinou, S. E. Prickett and M. L. Mavrovouniotis, “Estimation of Thermodynamic and Physical Properties of Acyclic Hydrocarbons using the ABC Approach and Conjugation Operators”, Ind, Eng. Chem. Res., 32, 1734 (1993).

[8] L. Constantinou, S. E. Prickett and M. L. Mavrovouniotis, “Estimation of Properties of Acyclic Organic Compounds using Conjugation Operators”, Ind, Eng. Chem. Res., 39, 395 (1994).

[9] S. E. Prickett, L. Constatinou and M. L. Mavrovouniotis, “Computational Identification of Conjugate Paths for Estimation of Properties of Organic Compounds”, Molecular Simulation, 11, 205 (1993).

[10] L. Constantinou and R. Gani, “New Group Contribution Method for Estimating Properties of Pure Compounds”, AIChE J., 40, 1697 (1994).

[11] J. Marrero-Morejón and E. Pardillo-Fontdevilla, “Estimation of Pure Compound Properties Using Group-Interaction Contributions”, AIChE J., 45, 615 (1999).

[12] CAPEC Database, Department of Chemical Engineering, DTU, Lyngby, Denmark (2000).

Table 1. Selected function for each property

Property (X)	Left-hand side of Eq. 1 [Function $f(X)$]	Right-hand side of Eq. 1 (Group-Contribution Terms)
Normal Melting Point (Tm)	$exp(Tm/Tm0)$	$\Sigma_i N_i Tm1i + \Sigma_j M_j Tm2j + \Sigma_k O_k Tm3k$
Normal Boiling Point (Tb)	$exp(Tb/Tb0)$	$\Sigma_i N_i Tb1i + \Sigma_j M_j Tb2j + \Sigma_k O_k Tb3k$
Critical Temperature (Tc)	$exp(Tc/Tc0)$	$\Sigma_i N_i Tc1i + \Sigma_j M_j Tc2j + \Sigma_k O_k Tc3k$
Critical Pressure (Pc)	$(Pc-Pc1)^{-0.5}-Pc2$	$\Sigma_i N_i Pc1i + \Sigma_j M_j Pc2j + \Sigma_k O_k Pc3k$
Critical Volume (Vc)	$Vc-Vc0$	$\Sigma_i N_i Vc1i + \Sigma_j M_j Vc2j + \Sigma_k O_k Vc3k$
Standard Gibbs Energy at 298 K (Gf)	$Gf-Gf0$	$\Sigma_i N_i Gf1i + \Sigma_j M_j Gf2j + \Sigma_k O_k Gf3k$
Standard Enthalpy of Formation at 298 K (Hf)	$Hf-Hf0$	$\Sigma_i N_i Hf1i + \Sigma_j M_j Hf2j + \Sigma_k O_k Hf3k$
Standard Enthalpy of Vaporization at 298 K (Hv)	$Hv-Hv0$	$\Sigma_i N_i Hv1i + \Sigma_j M_j Hv2j$
Standard Enthalpy of Fusion ($Hfus$)	$Hfus-Hfus0$	$\Sigma_i N_i Hfus1i + \Sigma_j M_j Hfus2j + \Sigma_k O_k Hfus3k$

Table 2. Values of the additional adjustable parameters

Adjustable Parameter (Universal Constants)	Value
Tm0	147.450 K
Tb0	222.543 K
Tc	231.239 K
Pc1	5.9827 bar
Pc2	0.108998 $bar^{-0.5}$
Vc0	7.95 cm^3/mol
Gf0	-34.967 kJ/mol
Hf0	5.549 kJ/mol
Hv0	11.733 kJ/mol
Hfus0	-2.806 kJ/mol

Table 3. Global comparison of consecutive first, second and third approximations

Property (*X*)	Data Points	STD			AAE			ARE (%)		
		1st	2nd	3rd	1st	2nd	3rd	1st	2nd	3rd
Tm (K)	1547	33.87	29.52	27.67	24.90	21.41	20.22	9.3	7.9	7.6
Tb (K)	1794	11.11	8.96	8.09	7.90	6.38	5.89	1.8	1.4	1.4
Tc (K)	783	17.25	8.50	6.99	8.75	5.67	4.93	1.4	0.9	0.8
Pc (bar)	775	1.73	1.53	1.39	1.02	0.87	0.79	2.9	2.6	2.3
Vc (cc/mol)	762	13.36	11.57	10.74	9.12	7.85	7.33	2.2	1.9	1.8
Gf (kJ/mol)	679	8.37	6.85	5.90	5.35	4.12	3.70	---	---	---
Hf (kJ/mol)	686	8.29	6.79	5.75	5.27	4.05	3.60	---	---	---
Hv (kJ/mol)	437	2.05	1.61	---	1.10	0.86	---	2.7	2.3	---
Hfus (kJ/mol)	711	4.16	3.88	3.65	2.58	2.32	2.17	18.3	16.4	15.7

$$\text{STD} = \sqrt{\frac{\sum \left(Xest - X\,\text{exp}\right)^2}{N}}$$

$$\text{AAE} = \frac{1}{N}\sum \left|Xest - X\,\text{exp}\right|$$

$$\text{ARE} = \frac{1}{N}\sum \left|\frac{Xest - X\,\text{exp}}{X\,\text{exp}}\right| 100\%$$

where N is the number of data points, $Xest$ is the estimated value of the property X, and $Xexp$ is the experimental value of the property X

Table 4. Comparison of average deviations for second and third order approximations

Property *(X)*	Data Points (total)	2nd			3rd		
		Data Points	Deviations 1st	Deviations 2nd	Data Points	Deviations 1st & 2nd	Deviations 3rd
Tm	1547	960	9.7 %	7.5 %	181	10.0 %	7.1 %
Tb	1794	1107	1.9 %	1.3 %	141	2.6 %	1.4 %
Tc	783	412	1.6 %	0.8 %	52	1.8 %	0.5 %
Pc	775	411	2.6 %	1.9 %	64	5.0 %	2.3 %
Vc	762	408	2.5 %	1.9 %	62	2.7 %	1.4 %
Gf	679	358	5.8 *	3.5 *	57	9.5 *	4.6 *
Hf	686	353	5.6 *	3.3 *	58	9.3 *	4.0 *
Hv	437	218	2.5 %	1.6 %	----	----	----
Hfus	711	351	19.2 %	15.5 %	99	22.9 %	17.7 %

Deviations are expressed as average relative errors for all properties excepting for *Gf* and *Hf*, which are expressed as average absolute errors

* kJ/mol

Table 5. Comparison of accuracy between a classical group-contribution scheme and the proposed method

	Data	STD		AAE		ARE (%)	
Property (*X*)	Points	JR	New	JR	New	JR	New
Tm (K)	1103	38.87	25.34	34.90	18.76	14.6	7.5
Tb (K)	1211	15.86	8.01	11.02	5.89	3.1	1.4
Tc (K)	587	18.73	6.87	10.96	4.87	2.1	0.9
Pc (bar)	573	3.71	1.36	2.45	0.74	5.6	2.2
Vc (cc/mol)	544	18.36	10.69	14.53	7.25	2.7	1.8
Gf (kJ/mol)	481	12.41	5.90	9.03	3.62	---	---
Hf (kJ/mol)	493	12.23	5.68	8.98	3.60	---	---
Hv (kJ/mol)	343	2.93	1.60	2.71	0.83	4.2	2.4
Hfus (kJ/mol)	499	6.84	3.62	3.06	2.11	46.3	15.6

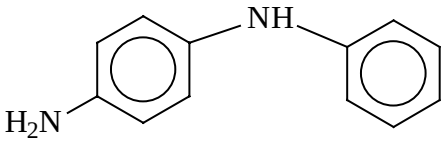
JR = Joback and Reid [1]

New = Proposed method

Appendix 1

To illustrate the proposed method, we provide the estimation of the normal boiling point and normal melting point using six example compounds. The experimental data and estimations through Joback and Reid method [1] are also given.

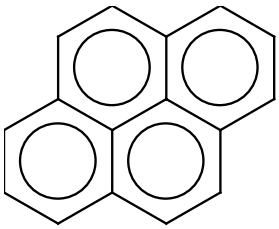
Example 1. Estimation of the normal boiling point of N-Phenyl-1,4-benzenediamine



(Experimental value: *Tb* = 627.15 K)

First-order Groups	Occurrences	Contribution
aC-NH2	1	3.8298x1
aC-NH	1	2.9230x1
aC	1	1.5468x1
aCH	9	0.8365x9
$\Sigma_i N_i Tb1i$ = 15.8281		
Tb = 222.543ln(15.8281) = 614.62 K		
(first-order approx., error: 12.53 K)		
Second-order Groups	Occurrences	Contribution
AROMRINGs1s4	1	0.1007x1
$\Sigma_j M_j Tb2j$ = 0.1007		
Tb = 222.543ln(15.8281+0.1007) = 616.03 K		
(second-order approx., error: 11.12 K)		
Third-order Groups	Occurrences	Contribution
aC-NH-aC	1	0.5768x1
$\Sigma_k O_k Tb3k$ = 0.5768		
Tb = 222.543ln(15.8281+0.1007+0.5768) = 623.94 K		
(third-order approx., error: 3.21 K)		
Estimation through Joback and Reid [1]: 655.20 K		
error: -28.05 K		

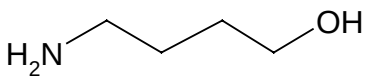
Example 2. Estimation of the normal boiling point of Pyrene



(Experimental value: *Tb* = 677.15)

First-order Groups	Occurrences	Contribution
aC (fused with arom. ring)	6	1.7324x6
aCH	10	0.8365x10
	$\Sigma_i N_i Tb1i$ = 18.7593	
Tb = 222.543ln(18.7593) = 652.43 K (first-order approx., error: 24.72 K)		
No second-order groups are involved		
Third-order Groups	Occurrences	Contribution
AROM.FUSED[3]	2	0.0402x2
AROM.FUSED[4p]	2	0.9126x2
	$\Sigma_k O_k Tb3k$ = 1.9056	
Tb = 222.543ln(18.7593+1.9056) = 673.96 K (third-order approx., error: 3.19 K)		
Estimation through Joback and Reid [1]: 651.56 K error: -24.41 K		

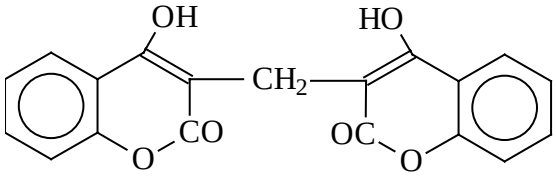
Example 3. Estimation of the normal boiling point of 4-aminobutanol



(Experimental value: *Tb* = 478.15 K)

First-order Groups	Occurrences	Contribution
OH	1	2.5670x1
CH2NH2	1	2.7987x1
CH2	3	0.7141x3
	$\Sigma_i N_i Tb1i$ = 7.508	
	Tb = 222.543ln(7.508) = 448.64 K	
	(first-order approx., error: 29.51 K)	
No second-order groups are involved		
Third-order Groups	Occurrences	Contribution
NH2-(CHn)m-OH (m>2)	1	1.0750x1
	$\Sigma_k O_k Tb3k$ = 1.0750	
	Tb = 222.543ln(7.508+1.0750) = 478.42 K	
	(third-order approx., error: -0.27 K)	
Estimation through Joback and Reid [1]: 364.31 K error: 113.84 K		

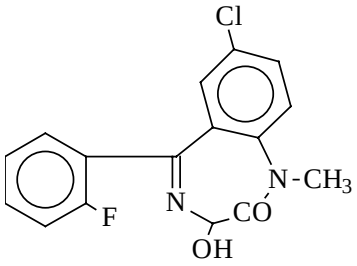
Example 4. Estimation of the normal melting point of 3,3'-Methylenebis-4-hydroxycoumarin (Dicoumarol)



(Experimental value: *Tm* = 563.15 K)

First-order Groups	Occurrences	Contrib.
OH	2	2.7888x2
aC (fused with non-arom. ring)	4	1.2065x4
aCH	8	0.5860x8
C=C (cyc)	2	0.3048x2
CO (cyc)	2	3.2119x2
O (cyc)	2	1.3828x2
CH2	1	0.2515x1
	$\Sigma_i N_i Tm1i$ = 25.1421	
Tm = 147.450ln(25.1421) = 475.46 K (first-order approx., error: 87.69 K)		
No second-order groups are involved		
Third-order Groups	Occurrences	Contrib.
AROM.FUSED[2]	2	0.2825x2
aC-(CHn=CHm)cyc (in fused rings)	2	0.2479x2
aC-O (cyc) (in fused rings)	2	-0.3545x2
(CHm=C)cyc-CHp-(C=CHn)cyc (in different rings)	1	16.8558x1
	$\Sigma_k O_k Tm3k$ = 17.2076	
Tm = 147.450ln(25.1421+17.2076) = 552.34 K (third-order approx., error: 10.81 K)		
Estimation through Joback and Reid [1]: 749.28 K error: -186.13 K		

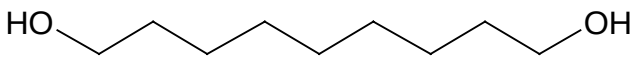
Example 5. Estimation of the normal melting point of 7-Chloro-5-(2-fluorophenyl)-1,3-dihydro-3-hydroxy-1-methyl-2H-1,4-benzodiazepin-2-one (Flutemazepan)



(Experimental value: *Tm* = 436.00)

First-order Groups	Occurrences	Contrib.
OH	1	2.7888x1
aC-Cl	1	1.7134x1
aC-F	1	0.9782x1
CH3	1	0.6953x1
aC	1	0.9176x1
aC (fused with non-arom. ring)	2	1.2065x2
aCH	7	0.5860x7
CH (cyc)	1	0.0335x1
CO (cyc)	1	3.2119x1
N (cyc)	1	0.6040x1
(C=N)cyc	1	6.6382x1
		$\Sigma_i N_i Tm1i$ = 24.0959
Tm = 147.450ln(25.1421) = 469.19 K (first-order approx., error: 33.19 K)		
Second-order Groups	Occurrences	Contrib.
AROMRINGs1s2	1	-0.6388x1
Ncyc-CH3	1	-0.0383x1
CHcyc-OH	1	1.3691x1
		$\Sigma_j M_j Tm2j$ = 0.6920
Tm = 147.450ln(24.0959+0.6920) = 473.36 K (second-order approx., error: 37.36 K)		
Third-order Groups	Occurrences	Contrib.
AROM.FUSED[2]s3	1	2.2589x1
aC-NHn(cyc) (in fused rings)	1	3.4983x1
aC-(C=N)cyc (in different rings)	1	-1.3060x1
aC-(CHn=N)cyc (in fused rings)	1	-10.1007x1
		$\Sigma_k O_k Tm3k$ = -5.6495
Tm = 147.450ln(24.0959+0.6920-5.6495) = 435.22 K (third-order approx., error: 0.78 K)		
Estimation through Joback and Reid [1]: 511.51 K error: -75.51 K		

Example 6. Estimation of the normal melting point of 1,9-Nonadiol



(Experimental value: *Tm* = 318.95 K)

First-order Groups	Occurrences	Contribution
OH	2	2.7888x2
CH2	9	0.2515x9
	$\Sigma_i N_i Tm1i$ = 7.8411	
Tm = 222.543ln(7.8411) = 303.66 K (first-order approx., error: 15.29 K)		
No second-order groups are involved		
Third-order Groups	Occurrences	Contribution
HO-(CHn)m-OH (m>2)	1	0.6674x1
	$\Sigma_k O_k Tm3k$ = 0.6674	
Tm = 222.543ln(7.8411+0.6674) = 315.70 K (third-order approx., error: 3.25 K)		
Estimation through Joback and Reid [1]: 312.83 K error: 6.12 K		